Plant Science 178 (2010) 523–530

Contents lists available at ScienceDirect

Plant Science

journal homepage: www.elsevier.com/locate/plantsci

Sequence variation at the rice blast resistance gene *Pi-km* locus: Implications for the development of allele specific markers☆

Stefano Costanzo, Yulin Jia*

USDA-ARS Dale Bumpers National Rice Research Center, Stuttgart, AR 72160, United States

ARTICLE INFO

Article history:
Received 15 January 2010
Accepted 16 February 2010
Available online 24 February 2010

Keywords:
Oryza sativa
Magnaporthe oryzae
Blast disease
Resistance gene
Molecular marker

ABSTRACT

The recently cloned blast resistance (*R*) gene *Pi-km* protects rice crops against specific races of the fungal pathogen *Magnaporthe oryzae* in a gene-for-gene manner. The use of blast *R* genes remains the most cost-effective method for an integrated disease management strategy. To facilitate rice breeding we developed a *Pi-km* specific DNA marker. For this purpose, we initially explored the existing sequence diversity for alleles of the two genes responsible for the *Pi-km* specificity. The analysis of 15 rice cultivars revealed that the majority of nucleotide polymorphisms were associated with the *Pi-km1* gene. Interestingly, the correspondent amino acid variation was localized within the predicted coiled-coil domain of the putative Pi-km1 protein. In contrast, the sequence of *Pi-km2* alleles was highly conserved even within distantly related cultivars. Furthermore, disease reactions of the selected cultivars to five *M. oryzae* isolates, as well as their determined *Pi-km1* allele, showed a good correlation with the known *Pi-k* genes (*-k*/*-kh*/*-km*/*-ks*/*-kp*) historically reported for these cultivars. Based on these findings, specific primer sets have been designed to discriminate among the various *Pi-km* alleles. The new markers should simplify the introgression of the valuable blast resistance associated with the complex *Pi-k* locus into rice cultivars.

Published by Elsevier Ltd.

1. Introduction

Rice is one of the most important staple foods in the world [1], and rice blast caused by the ascomycete fungal pathogen *Magnaporthe oryzae* B. Couch (formerly *Magnaporthe grisea*), still represents a major constraint for worldwide rice production. The use of host resistance remains the most cost-effective method for disease management strategy [2]. Over the years, a number of resistance (*R*) genes have been identified in diverse rice germplasm that recognize pathogen elicitors and trigger a defense mechanism ultimately leading to resistance. The classic gene-for-gene hypothesis proposed a specific interaction involving the products of a *R* gene in the host and of an avirulence (*AVR*) gene in the pathogen [3]. The aggressiveness of *M. oryzae* isolates is determined by the frequent occurrence of virulent races able to defeat the effectiveness of deployed *R* genes [4]. To date, more than 80 blast *R* genes have been identified [5], and twelve of them have also been molecularly characterized: *Pi-b* [6], *Pi-ta* [7], *Pi-9* [8], *Pi-2*/*Pi-zt* [9], *Pi-d2* [10], *Pi-36* [11], *Pi-37* [12], *Pi-km* [13], *Pi-5* [14], *Pi-t* [15], *Pi-d3* [16] and *pi-21* [17]. The fine mapping and cloning of the *Pi-kh* gene from 'Tetep' has also been reported by Sharma et al. [18], however lack of supporting evidence casts significant doubts on this claim [19]. In plants, the most prevalent class of *R* genes is composed of proteins characterized by a centrally located nucleotide-binding site (NBS) region associated with a C-terminal leucine-rich repeat (LRR) domain. The LRR region is generally thought to be involved in recognition of the pathogen effector proteins. Within the rice genome, more than 600 *R* genes that belong to the NBS-LRR class have been computationally identified [20].

Molecular markers tightly linked to major *R* genes represent an important tool for marker-assisted selection (MAS) especially when they can be applied during the early phase of plant growth and when they are used to circumvent the association with undesirable agronomical traits (linkage drag). Despite the growing number of blast *R* genes that have been fine mapped and/or cloned, there are only few examples where this new knowledge has had a direct impact in the breeding process. Among them, DNA markers derived from two cloned blast *R* genes *Pi-b* and *Pi-ta* have been widely implemented and are currently used in several rice breeding programs [21], as well as a PCR-based SNP markers for genes at the *Piz* locus [22]. Conventional breeding with MAS would therefore benefit from the development of new *R* gene specific markers, which would allow pyramiding multiple genes

☆ Sequence data from this article have been deposited with the EMBL/GenBank Data Libraries under accession nos. GU811849-GU811872.

* Corresponding author.
E-mail address: yulin.jia@ars.usda.gov (Y. Jia).

0168-9452/$ – see front matter. Published by Elsevier Ltd.
doi:10.1016/j.plantsci.2010.02.014

in adapted germplasm to achieve a broader spectrum of disease resistance.

The blast *R* gene *Pi-km* [13] was recently fine mapped and cloned from the Japonica rice cultivar 'Tsuyuake' through a map-based cloning strategy. Results of complementation analysis showed that the *Pi-km* specificity is functional only with the concurrent expression of two adjacent NBS-LRR class genes: *Pikm1-TS* and *Pikm2-TS*. The *Pi-km* genes are located in proximity of the telomeric region of the long arm of chromosome 11. This chromosomal area is known to harbor a large number of blast *R* genes including *Pi-44*(*t*) [23] and *Pi-1* [24], and several bacterial blight *R* genes: *Xa2*, *Xa4*, *Xa22*(t) and *Xa26* [25–27]. *Pi-km* represents the first cloned member of this dense *Pi-k* cluster, which includes other tightly linked or allelic blast resistance genes such as *Pi-k*, *Pi-kh*, *Pi-kp* and *Pi-ks* [19,28,29]. Several studies have previously attempted to fine map and elucidate the individual contribution of resistance derived from this chromosomal region, but the high density of *R* genes analogues has proven to be a major obstacle towards this goal. A renewed interest in this research was brought by two recent studies: The first, by Yoshida et al. [30], where the corresponding pathogen-secreted virulence protein AVR-Pik/km/kp was identified through gene association analysis in a *M. oryzae* field isolate Ina168. The second, by Ashikawa et al. [31], provided an initial attempt of characterizing the allele diversity of the *Pi-km* locus by determining and comparing the sequence of their corresponding LRR regions.

The objectives of the present study were to survey the complete sequence variation for the two *Pi-km* genes in selected genetically distinct U.S. rice cultivars, and subsequently develop closely linked DNA markers for this *R* gene that could be easily incorporated into current MAS breeding programs. Cloning and molecular characterization of *R* genes can lead to a better understanding of molecular mechanisms of *R* gene meditated resistance and also benefit the development of new rice cultivars using DNA markers specifically targeted to unique regions of the cloned genes. In this article, we report the sequence analysis for the *Pi-km* locus and the development and testing of a series of dominant and co-dominant PCR based markers to assist in the detection of *Pi-km* in rice cultivars. Based on sequence analysis and computational translation we report the existence of at least six alleles within this locus. Furthermore, artificial blast inoculations revealed different resistance spectra to distinct blast isolates for some of the rice cultivars tested, which are historically associated with particular *Pi-k* genes. The occurrence of high levels of sequence polymorphism, determined in this study, brings new insight into the complex nature of this *R* gene locus, which may impact the recognition of the pathogen-encoded elicitors initiating signaling pathways leading to a defense response.

2. Materials and methods

2.1. Plant material

Seeds of the rice cultivar 'Katy' (PI 527707) [32], were kindly provided by Dr. Karen Moldenhauer (University of Arkansas, Rice Research and Extension Center, Stuttgart, AR). Seeds of rice cultivars 'Shin 2', 'Kanto 51', 'Pusur' and 'Toto' were provided by Dr. H. Bockelman from the USDA-ARS, National Small Grains Germplasm Research Facility, Aberdeen, Idaho. All other accession seeds were provided by the Genetic Stocks – *Oryza* (GSOR) Collection at the USDA-ARS Dale Bumpers National Rice Research Center, Stuttgart, AR. A complete list of the cultivars included in the present work is shown in Table 3 together with their correspondent accession identifier numbers. For seven of the 28 cultivars we have also included their documented association with a specific *Pi-k* genes (*-k*/*-kh*/*-km*/*-ks*/*-kp*). Seeds were germinated and plants were grown under greenhouse conditions at 24–30 °C with a light and dark cycle of 16 and 8 h, respectively. Plants grown in a greenhouse for approximately 2–4 weeks (four-leaf seedling stage) were used for pathogenicity evaluations and genomic DNA isolation.

2.2. Fungal isolate, culture and infection assay

All five *M. oryzae* isolates used in this study were obtained from the collection maintained at the USDA Dale Bumpers National Rice Research Center, in Stuttgart AR. The *M. oryzae* race IB54 (unnamed isolate, UI), IB45 (UI), IC17 (ZN60), IG1 (ZN39) and IH1(UI) were all collected in the southern United States and routinely cultured on oatmeal agar medium. Pathogen inoculation assays were carried out as previously described by Valent et al. [33]. Conidial concentration was measured with a hemacytometer, and the final concentration was adjusted to 1.4×10^6 spores/mL. Rice plants at the three to four-leaf stages were placed in large plastic bags and inoculated with approximately 15 mL of conidial suspension using an airbrush. After inoculation, plastic bags were sealed to

Table 1
Primer pairs utilized for dominant, co-dominant markers and SNuPE assay.

Gene[a]	Primer designation[a]	Primer sequence (5′ → 3′)	Annealing temperature (°C)	Marker feature[b]	Expected fragment/s size (bp)
Pi-km1	Ckm1F	TGAGCTCAAGGCAAGAGTTGAGGA	56	Co-dominant	174/213
	Ckm1R	TGTTCCAGCAACTCGATGAG			
Pi-km2	Ckm2F	CAGTAGCTGTGTCTCAGAACTATG	60	Co-dominant	290/332
	Ckm2R	AAGGTACCTCTTTTCGGCCAG			
Pi-km1	Dkm1F	CTGGAGAGCTTCCGTGTCGAC	60	Dominant	223
	Dkm1R	TCTTCACGACGTCAATGGTGGC			
Pi-km2	Dkm2F	GTTGTTCACTCCGTATCTACTACGTC	60	Dominant	291
	Dkm2R	TTCCTCCGTGATCTCAGCAACG			
Pi-km1	KMsnupeF	ATCAAGTTGCTGGAACAAGG	56	Single-nucleotide primer extension (SNuPE) assay	496/524/534
	KMsnupeR	TTGACCTCGTGAAATTCACA			
	KM-A	TCGYMGGTGAYCTAAGAGAC		G/C/A	21
	KM-B	CAATCGCCGGTGACCTAAGAGACGA		T/G	26
	KM-C	CYGCGCTCCGGAAGAAGGTGGGC		C/G/T	24
	KM-D	GRTCTCYGCGCTCCGGAAGAAGGTGGGCC		A/C	30

[a] Specific primers were designed to amplify the two distinct genes (Pi-km1 and Pi-km2) involved in Pi-km resistance.
[b] The SNuPE assay was developed based on the sequence variation determined exclusively in Pi-km 1 alleles.

maintain a high level of humidity and incubated at 25 °C in low light conditions for 24 h. After the initial incubation period, plants were removed from the bags and returned to the greenhouse. A final visual evaluation of treated plants was recorded 7 days post-inoculation to establish plant disease reaction.

2.3. Nucleic acid isolation

Rice genomic DNA was isolated from leaf samples that were harvested and immediately frozen in liquid nitrogen and stored at −80 °C until processed. DNA was isolated using DNeasy Plant Mini Kit (Qiagen, Valencia, CA) from 0.1 g of leaf tissue following the manufacturer's instructions.

2.4. PCR amplification and sequencing

The DNA sequence of the *O. sativa* cv. Tsuyuake BAC clone TS18H12 (GenBank accession no. AB462256), was used as reference sequence for primer design. PCR primers were designed using the Primer-BLAST tool available from the NCBI web site that is based on Primer3 software (http://blast.ncbi.nlm.nih.gov/Blast.cgi) [34]. Primer pairs were designed to amplify approximately 1.5 kb overlapping DNA fragments that span the 5′ promoter region to the 3′ termination area for each of the two genes. These primer sets were subsequently utilized to amplify and sequence the corresponding fragments in 12 selected rice cultivars. Templates PCR reactions were set up as follows: 1× PCR buffer, 0.2 mM of each dNTP, 1.5 mM of $MgCl_2$, 0.5 μM each primer, 50 ng of genomic DNA template, 1.5 units of TaKaRa LA *Taq* DNA Polymerase (Takara Bio Inc., Shuzo, Kyoto, Japan), and PCR grade water to final volume of 20 μL. After amplification, all PCR products were verified for size and integrity by electrophoresis on 1% agarose gels in TAE buffer following standard procedures [35]. PCR products were purified using QIAquick PCR Purification Kit (QIAGEN Valencia, CA) according to manufacturer's protocol. After purification, all DNA templates were quantified using an ND-1000 NanoDrop spectrophotometer (NanoDrop Technologies, LLC, Wilmington, DE) and then submitted to the USDA Mid South Area Genomics Laboratory (MSAGL) in Stoneville MS for sequencing.

2.5. Sequence scans and analysis

All individual raw DNA sequences were assembled into contigs using SeqMan and aligned with the MegAlign implementing the Clustal W method available from the DNASTAR Lasergene 6.1 software package (DNASTAR Inc., Madison, WI). Pairwise comparisons between the Tsuyuake Pi-km1 and Pi-km2 protein sequences and their respective homologues in all other cultivars were performed using the Matcher program (http://mobyle.pasteur.fr/cgi-bin/portal.py?form=matcher). The scoring matrix file used when comparing protein sequences was BLOSUM62, with a gap penalty value of 11 and a gap extension penalty of 1. Phylogenetic and molecular evolutionary analyses were conducted using MEGA version 4.0 [36], applying the Neighbor-Joining algorithms. The phylogenetic trees were constructed using the neighbor-joining (NJ) method and drawn using the MEGA program. The stability of each tree was evaluated by bootstrap analysis with 1000 replications. Prediction of coiled-coil domains in protein sequences were identified using the 14/28-residue window output from the PCOILS program and the MTIDK profile matrix [37] (http://toolkit.tuebingen.mpg.de/pcoils). Position-specific variability in amino acid sequence and the overall pattern of conservation within whole protein sequences was analyzed using the Plotcon program (http://emboss.sourceforge.net/) in the EMBOSS package [38].

3. Results

3.1. Sequence analyses

The sequences of a 12 kb genomic region matching the cv. Tsuyake *Pi-km* locus on chromosome 11 (obtained from GenBank accession no. AB462256), were initially identified through BLAST homology searches to cultivars 'Nipponbare' and '93-11' using the Gramene (http://www.gramene.org/) database websites. Using the blastp (protein-protein BLAST) algorithm, two putative homologue sequences of Pi-km1-TS and Pi-km2-TS were identified in both genomes. In Nipponbare, two adjacent sequences in proximity of the telomeric region of chromosome 11 had 59.2% (Pi-km1) and 76.9% (Pi-km2) amino acid identity values (Table S.1). In contrast, the corresponding sequences in the genome of 93-11 were located on the short arm of chromosome 2 with 57.9% (Pi-km1) and 77.0% (Pi-km2) identity values. To identify polymorphic regions that could be exploited for marker development, we compared their sequences by ClustalW multiple alignment. Based on the alignment results of the individual *Pi-km1* and *Pi-km2* homologues sequences, we initially designed one set of dominant and co-dominant PCR based markers for both genes. The list of primer sequences for these markers is shown in Table 1. The dominant markers, referred to as Dkm1 and Dkm2, were specifically designed to amplify a fragment of 223 bp (*Pi-km1*) and 291 bp (*Pi-km2*) when using Tsuyuake genomic DNA and would not amplify any product from Nipponbare and 93-11genomic DNA. The co-dominant markers, Ckm1 and Ckm2, would amplify a 174 bp (*Pi-km1*) and 290 bp (*Pi-km2*) from Tsuyuake genomic DNA and 213 bp (*Pi-km1*) and 332 bp (*Pi-km2*) from both Nipponbare and 93-11. We tested these markers on 22 rice cultivars for which we had available genomic DNA samples. Results of the PCR amplifications with the markers were reported on Table 2. The results from the dominant markers identified only two groups based on their concurrent amplification ("Tsuyuake type") or non-amplification ("Nipponbare and 93-11 type") of the 223 bp and 291 bp fragments. In contrast, the PCR amplification with co-dominant markers recognized four types: 174–290 bp (Tsuyuake type); 213–332 bp (Nipponbare and 93-11 type); no-amplification (NA) with Ckm1-290 bp, and 213–290 bp.

To further investigate the observed DNA sequence variability within homologues of two genes at the *Pi-km* locus, a 12 kb region, corresponding to the cv. Tsuyake *Pi-km* locus, was PCR-amplified, sequenced and assembled for 12 selected rice cultivars. The complete sequence of this locus was obtained for three additional cultivars (Nipponbare, 93-11 and Tsuyuake) from the NCBI and Gramene databases. The predicted introns for both genes in all of the selected cultivar were identified based on sequence conservation of the exon-intron boundaries with the corresponding Tsuyuake sequence. After intron removal, the two open reading frames (ORFs) with high homology to the protein sequences of Pi-km1-TS and Pi-km2-TS as reported by Ashikawa et al. [13] were obtained. For both genes, the predicted CDSs were verified using BioEdit software [39]. Based on those predictions, each individual copy was translated into amino acid sequences and aligned for comparisons.

The protein sequence analysis revealed that Pi-km1 homologues varied in length from 778 to 1145 amino acid residues, with five distinct protein sizes present and the majority of cultivars (9 cvs.) coding for a protein of 1142 aa. For Pi-km2, protein size varied between 1021 (13 cvs.) and 1044 aa. For Pi-km1, the amino acid sequence identity values obtained ranged, for the majority of them, between 94.8 and 100.0%. Only cultivars Nipponbare and 93-11 had the lower values (59.2 and 57.9%, respectively). When comparing Pi-km2 homologues, amino acid identities ranged from 99.6 to 100.0%, for the uniform group, and had lower amino acid identities values (77.0 and 76.9%) for cultivars 93-11 and Nipponbare,

Table 2
Test of dominant and co-dominant markers on rice cultivars and breeding lines for presence of *Pi-km1* and *Pi-km2*.

Cultivar[a]	Accession identifier	Dkm1[b]	Dkm2	Pi-km results	Ckm1	Ckm2	Pi-km[c] results
Leah	GSOR 310045	223	291	+/+	174	290	+/+
Guang Lu Ai 4	PI596840	223	291	+/+	NA	290	−/+
IR64	PI497682	223	291	+/+	174	290	+/+
Te Tep	PI389176	223	291	+/+	174	290	+/+
Kitake	PI652747	223	291	+/+	174	290	+/+
Tsuyuake	PI597057	223	291	+/+	174	290	+/+
Lemont	PI475833	223	291	+/+	174	290	+/+
Cypress	PI561734	223	291	+/+	174	290	+/+
RU9101001	PI 651498 MAP	223	291	+/+	174	290	+/+
Dawn	PI312773	223	291	+/+	213	290	−/+
Wells	PI 612439	223	291	+/+	174	290	+/+
M202	PI494105	NA	NA	−/−	213	332	−/−
Katy/M2354	PI527707	223	291	+/+	174	290	+/+
Reiho	PI403943	223	291	+/+	174	290	+/+
Caloro	PI 388616	223	291	+/+	174	290	+/+
Drew	PI 596758	223	291	+/+	174	290	+/+
IR36	PI408586	223	291	+/+	174	290	+/+
93-11	–	NA	NA	−/−	213	332	−/−
Nipponbare	PI514663	NA	NA	−/−	213	332	−/−
Lagrue	PI568891	NA	NA	−/−	213	332	−/−
Amane	PI373335	NA	NA	−/−	213	332	−/−
C101PKT	PI602663	NA	NA	−/−	213	332	−/−

[a] Includes a breeding line (RU9101001) and M2354 (a mutant line derived from Katy).
[b] NA, no amplification using dominant markers Dkm1 and Dkm2.
[c] The symbol (−) for co-dominant markers indicates no amplification or amplification of the 'non-Tsuyuake' type.

respectively. For the 15 cultivars selected in this study, the analysis of nucleotide sequence conservation among the putative coding regions of the two genes, revealed a distinctive pattern. While a high level of residue conservation throughout the entire amino acid sequence characterized *Pi-km2* homologues, *Pi-km1* homologues showed a selective accumulation of polymorphism in the first third of their CDS, corresponding to the sequence between 400 and 1000 bp from the start codon (Fig. 1). The remaining regions of the *Pi-km1* CDS showed levels of sequence conservation comparable to those found for *Pi-km2*.

3.2. Phylogenetic analyses

For the selected 15 rice cultivars, separate amino acid sequences alignments for Pi-km1 and Pikm2 were conducted with ClustalW and the Neighbor-Joining method for phylogenetic inference. The three reconstructions showed a distinct clustering of sequences in three main similarity groups for the Pi-km1 homologues, while only two major groups were present for Pi-km2 sequences (Fig. 2). For both proteins analyzed, the Nipponbare and 93-11 sequences clustered together forming a distinct out-group separated from the

Table 3
Correlation between reported *Pi-k* genes for rice cultivars and breeding lines and amino acid type identified at Pi-km1 aa position I and II.

Cultivar[a]	Accession identifier	Reported Pi-k gene	aa position I/II	% positives[b]	Reference
Pusur	PI406083	-kp	K/D	96	Kiyosawa [46]
Dular	PI403392		K/D	96	
Leah	GSOR 310045		K/D	96	
Toto	PI239074	-k	K/H	96	Kiyosawa [47]
Guang Lu Ai 4	PI596840		K/H	–	
IR64	PI497682		K/H	–	
Te-tep	PI389176	-kh	D/P	99	Kiyosawa [48]
Kanto 51	PI389183	-k	E/H	99	Yamasaki et al. [49]
Kitake	PI652747		E/H	–	
Tsuyuake	PI597057	-km	Q/P	100	Ashikawa et al. [13]
Lemont	PI475833		Q/P	100	
Cypress	PI561734		Q/P	100	
RU9101001	PI 651498 MAP		Q/P	100	
Dawn	PI312773		Q/P	–	
Wells	PI 612439		Q/P	–	
M202	PI494105		R/Y	–	
Shin2	PI431054	-ks	E/P	99	Kiyosawa et al. [50]
Katy/M2354	PI527707	-ks	E/P	99	Jia et al. [44]
Reiho	PI403943		E/P	99	
Caloro	PI 388616		E/P	–	
Drew	PI 596758		E/P	–	
Bengal	PI561735		E/P	–	
IR36	PI408586		E/P	–	
93-11	-		D/H	68	
Nipponbare	PI514663		D/R	68	
LaGrue	PI568891		D/R	–	
Amane	PI373335		D/R	–	
C101PKT	PI602663		D/R	–	

[a] Includes a breeding line (RU9101001) and M2354 (a mutant line derived from Katy).
[b] The value of % identity match is reported only for cultivars where a complete sequence of Pikm1 was obtained.

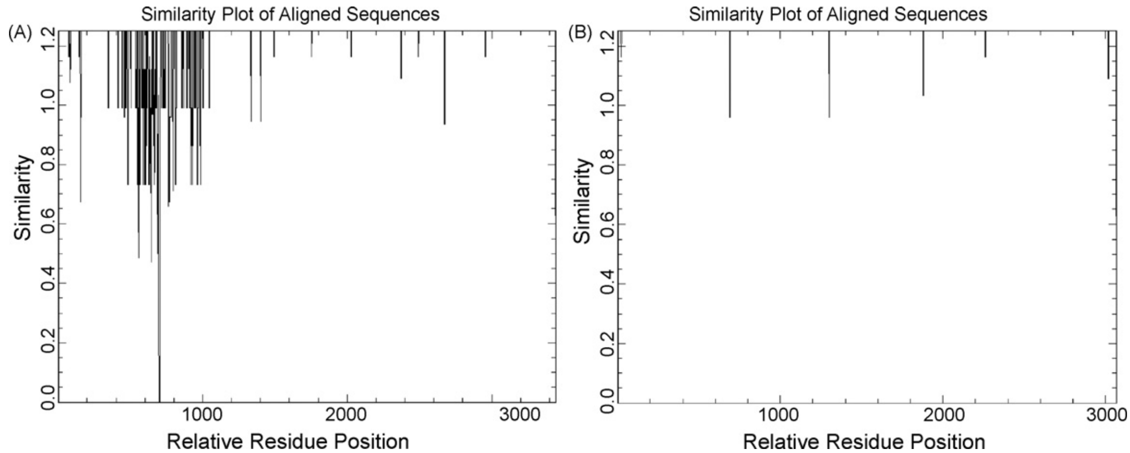

Fig. 1. Conservation plots of the CDS sequence alignment of 13 Pi-km1 (A) and Pi-km2 (B) alleles. Sequence from cultivars Nipponbare and 93-11 were excluded from the conservation analysis.

remaining cultivars. In the Pi-km1 phylogenetic tree, the sequences of cvs. Toto, Leah, Dular and Pusur formed a sub-group, appearing more divergent from the remaining sequences. This separation was not maintained for the Pi-km2 sequences, which, as previously mentioned, appeared much more conserved and uniform within these selected genotypes.

3.3. SNuPE assay development

Based on this result, we focused solely on Pi-km1 homologues and expanded the initial analysis with 13 additional cultivars, focusing the analysis on a 600 bp sequence that represented the most polymorphic region of the gene. A complete list of the rice cultivars analyzed for this purpose is shown in Table 3. The obtained sequences were manually aligned to the initial 15 complete *Pi-km1* homologue sequences and translated into amino acid sequences. From the analysis of the multiple sequence alignment, two highly variable positions were identified corresponding to Pikm1-TS aa residues 229-Q and 252-P (indicated by 'I' and 'II' in Fig. 3). The combination of the two amino acid residues identified in positions I and II was found to be an efficient "parameter" that was used to characterize the observed variability within this individual gene. As shown in Table 3, the 28 different cultivars included in this study were classified in 9 distinct sub-groups based on their combined aa residues from position I and II. A method that has been shown to allow accurate discrimination of DNA sequence variants is the single-nucleotide primer extension assays (SNuPE) [40]. For this assay, a target region is amplified by PCR followed by a single base sequencing reaction that involves the use of primers annealing one base short of the polymorphic sites and implementing fluorescently labeled dideoxynucleotide terminators that interrupts the DNA sequence extension. The obtained products are subsequently run through a capillary sequencer for detection. For this purpose, we designed a set of primers (KMsnupe) to PCR amplify a product of variable size depending on the specific genotype (496/524/534 bp) that could be used in a SNuPE assay in combination with specific primers such as KM-A, KM-B (for two polymorphic sites in position I) and KM-C, KM-D (for two polymorphic sites in position II) (Table 3). Specifically, KM-A would be used to determine the first nucleotide within polymorphic site I (G, C or A), while KM-B would discriminate between T/G in the third position of the codon. The second polymorphic site would require the use of primer KM-C to discriminate C/G in first position, and primer KM-D for A/G/C in the second position within the codon. The simultaneous use in a single assay of all four KMsnupe primers would distinguish six out of the nine Pi-km1 alleles identified in the present study (Table 3), with the exception of those described for Nipponbare, 93-11 as well as cv. M202.

3.4. Blast inoculation tests

The U.S. collected blast isolates of races IB54, IB45, IC17, IG1 and IH1 were all tested in greenhouse experiments on 16 cultivars with detailed sequences of the *Pi-km* locus to establish their dis-

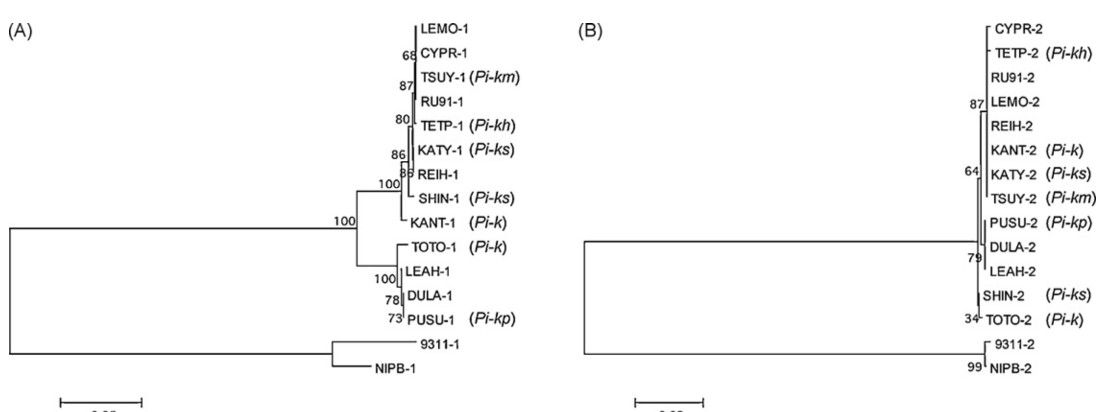

Fig. 2. Unrooted NJ phylogenetic tree based on aa sequences of the Pi-km1 (A) and Pi-km2 (B) alleles determined using the MEGA (v.4.0) software package. Numbers at the branch nodes represent bootstrap values as the proportions of 1,000 replications. Scale bar represent weighted sequence divergence. Reported Pi-k genes are indicated in parentheses following the abbreviated cultivar name. Full names in alphabetical order are: 93-11; Cypress; Dular; Kanto 51; Katy; Leah; Lemont; Nipponbare; Pusur; Reiho; RU9101001; Shin 2; Tetep; Toto; Tsuyuake.

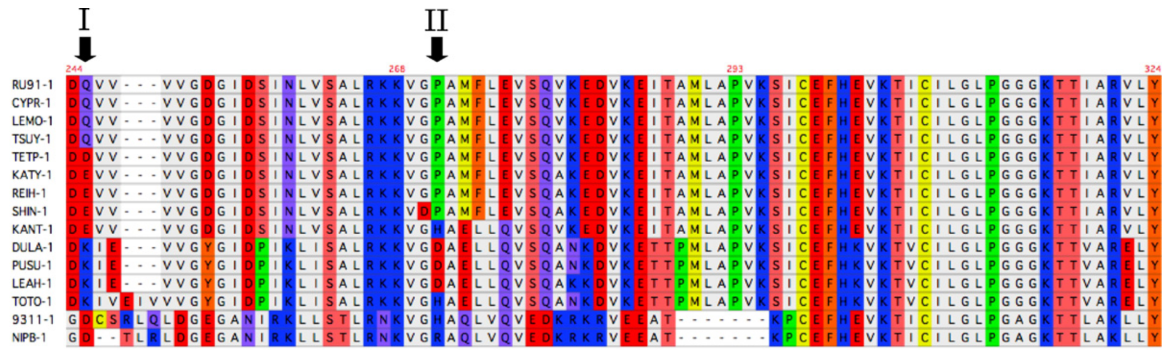

Fig. 3. A section of 80 aa of the protein ClustalW alignment of Pi-km1 homologues in15 rice cvs. Indicated by arrows are the two positions with variable aa residues targeted for allele specific marker development.

Table 4
Disease reactions of selected rice cultivars and breeding lines to five races of *Magnaporthe oryzae*.

Cultivar[a]	Reported Pi-k gene	IB-45	IB-54	IC-17	IG-1	IH-1
Tsuyuake	Pi-km	R	R	S	R	R
Lemont		R	R	S	R	R
Cypress		R	R	S	R	R
RU9101001		R	R	S	R	R
Shin2	Pi-ks	S	R	S	S	S
Katy	Pi-ks	R	R	R	R	R
M2354		S	R	S	S	S
Reiho		R	R	R	R	R
Pusur	Pi-kp	R	S	S	R	R
Dular		R	S	S	R	R
Leah		R	S	S	R	R
Toto	Pi-k	S	I	S	S	S
Tetep	Pi-kh	R	R	R	R	R
Kanto 51	Pi-k	R	R	S	S	R
93-11		S	S	S	S	S
Nipponbare		S	S	S	S	S

R, resistant; S, susceptible; I, intermediate.
[a] Includes a breeding line (RU9101001) and M2354 (a mutant line derived from Katy).

ease reaction. As shown in Table 4, only cvs. Katy, Reiho and Tetep appeared resistant to all five *M. oryzae* races tested. In contrast, cvs. Nipponbare, 93-11 and Toto, had a susceptible reaction to all isolates, with the sole exception of Toto that showed an intermediate reaction against race IB54. Of particular interest is the reaction of a fast neutron mutant line derived from the cv. Katy: M2354. This line displayed a complete susceptibility reaction except for race IB54. The same resistance spectrum was also shared by cv. Shin 2. Four isolates, Tsuyuake, Lemont, Cypress and the breeding line RU9101001, were all resistant to four of the five races, while cvs. Pusur, Dular and Leah were susceptible to both IB54 and IC17. A unique resistance spectrum was observed for cv. Kanto 51, which was only susceptible to races IC17 and IG1.

4. Discussion

Although two genes have been found necessary to confer *Pi-km* specific blast resistance, the higher level of polymorphism detected in the present study exclusively associated with the CDS of *Pi-km1* may suggest a more complex role for this gene in the host defense mechanisms. Support of this hypothesis comes from the results of the analysis with the PCOIL program. Interestingly, this region partially coincides with the predicted coiled-coil domain of this NBS-LRR gene, which has a 70% probability for residues 74–87, and 60% for residues 152–165, with a window size of 14, and a 61% probability for residues 71–98, with a window size of 28. The coiled-coil domain of NBS-LRR proteins has been suggested to take an active role in a plethora of protein-protein interactions and recognitions [41]. Noteworthy is the fact that, under the same conditions, a CC domain for the second NBS-LRR protein Pi-km2 was not statistically supported. It is possible that the observed variation in this protein region is the result of variable effector proteins interacting with the various alleles of this *R* gene. In contrast, the high level of sequence conservation within Pi-km2 homologues would suggest a more fundamental role for this second NBS-LRR gene similar to the tomato Prf protein [42]. A comparative study of other known cases in which two genes are required to confer resistance, such as the rice *Pi-5* [14] or the tomato *Pto*/*Prf* [42], may provide a better understanding of the individual gene specialized function. Surprisingly, the LRR regions encoded by both *Pi-km1* and *Pi-km2* genes were instead highly conserved. This observation was also confirmed in a recent publication by Ashikawa et al. [31] indicating a low Ka/Ks ratios in the LRR region of the two *Pi-km* genes in 16 elite rice cultivars and 35 landraces with a diverse geographic origin. Generally, a higher level of polymorphism is expected in the LRR region, which is thought to be involved in the recognition of effector proteins, and consequently the result of the evolutionary pressure by virulent pathogens races on the host [43].

A significant sequence variation separated the two cultivars Nipponbare and 93-11 from all others analyzed. This observation remains true for both *Pi-km1* and *Pi-km2* sequences. Although, as previously mentioned, the overall rate of divergence among *Pi-km1* homologues is more pronounced than those for *Pi-km2*. In particular, a major difference was found in the relative size of the first and second intron of *Pi-km1* in cvs. Nipponbare and 93-11 relative to those reported for the cv. Tsuyuake and their corresponding predicted sizes in all other cultivars analyzed. Furthermore, a major genome reorganization of the equivalent region on chromosome 11 between cv. Nipponbare and Tsuyuake, has been also reported by Ashikawa et al. [13] (Supplementary data Fig. S1). Our preliminary data suggests that a similar genome organization might be shared between cv. Tsuyuake and Katy. For this purpose, four different sets of primers were designed to selectively amplify a (~1.5 kb) region of four predicted *R* genes (LOC_Os11g46070; 46080; 46100 and 46130), located within the 127 kb region present in the Nipponbare genome but lacking in Tsuyuake (Fig. S1). These primer sets PCR amplified the expected fragment only from Nipponbare genomic DNA samples and failed when using Katy genomic DNA samples. In contrast, we were able to PCR amplify from Katy DNA samples but not for Nipponbare, four randomly selected fragments located within the 70 kb region present in cv. Tsuyuake but missing in Nipponbare. Katy is a U.S. tropical japonica rice cultivar that, in the southern U.S.A., is resistant to predominant *M. oryzae* races and it is still being used in several blast resistance breeding programs. Katy has been previously shown to carry several blast *R* genes including *Pi-ta* and possibly *Pi-ks* [44]. Pathogenicity tests on Katy have shown that *Pi-ta* confers resistance to races IB1, IB45, IB49, IC17, IH1, IE1 and IG1 while *Pi-ks* was believed to be responsible for the resistance to IB54. *Pi-ta* in Katy was originally introgressed from a landrace indica variety Tetep from Vietnam [29]. The Katy-derived fast neutron mutant M2354 is defective for *Pi-ta*-mediated disease response, but the induced mutation appears not to have affected the recognition specificity provided by the presence of *Pi-ks* [45]. The observed resistance spectra, in the present study, of the cv. Shin 2, historically associated with the *Pi-ks* resistance specificity, would also support the presence of *Pi-ks* in Katy and consequently in M2354.

It should be noted that the genomic sequence of the *Pi-km* locus for the indica cultivar 93-11 as reported in the Gramene database http://www.gramene.org/, is incomplete towards the 3′ end of the *Pi-km1* homologue. It is probable that the absence of sequence data in this 3′ end region has affected the accuracy of the gene prediction model for this gene. Furthermore, in the current genome assembly of 93-11 from the BGI Rise Database (http://rise2.genomics.org.cn/page/rice/index.jsp), the best matches for homologues of the two *Pi-km* genes are mapped on the short arm of chromosome 2 (2:7,152,000–7,162,000). To be noted that BLAST homology searches of the *O. sativa* (indica cultivar-group) WGS contigs database from NCBI (http://www.ncbi.nlm.nih.gov), identified two contigs (GenBank accession nos. AAAA02039802 and AAAA02047283), which are almost identical to respective portions of Pikm1-TS and Pikm2-TS, but have not been assigned a chromosomal location in the current genome assembly. This observation would indicate the existence in 93-11 of other homologues of the two *Pi-km*-TS genes. In addition, while a full length cDNA (FLcDNA) is available in support of the gene model of *Pi-km2*, in the cv. Nipponbare (Fig. S2), the FLcDNA partially associated with the *Pi-km1* sequence is from a transcript (GenBank accession no. AK067598) corresponding to the LOC_Os11g46220.1 sequence (annotated as hypothetical protein), mapped in proximity of the 3′ end of the *Pi-km2* homologue.

In conclusion, the set of dominant and co-dominant markers developed in the present work, were mainly able to distinguish between the "Tsuyuake-type" *Pi-km1* and *Pi-km2* alleles from the more distantly related homologues found in Nipponbare and 93-11 genomes. Based on the results of the sequence analysis that included a larger set of rice genotypes, we subsequently designed a more discriminative test (SNuPE assay) that would allow the identification of six of the nine *Pi-km1* alleles identified. A good correlation was also established between the obtained blast resistance spectra and the particular *Pi-km* alleles carried by the 16 cultivars tested, especially in respect of the previously reported *Pi-k* genes (*-k*/*-kh*/*-km*/*-ks*/*-kp*) for some of these cultivars. The development of these new markers provides a useful tool for tagging the distinct *Pi-k* alleles in conventional rice breeding using MAS. Furthermore, the sequences of the *Pi-km* locus obtained in the present work may contribute towards the understanding of the resistance specificity associated with this locus. However, as for other specific primer sets developed to date, suitability of primers used in this study have to be tested further with a broader rice genotype sample and testing various *M. oryzae* isolates including known differential lines.

Acknowledgements

The authors thank Michael Lin for assistance in plant artificial inoculations and greenhouse operations, Kristen Pratt for technical assistance, Seonghee Lee and Jackson Aaron for useful discussions and comments. We also thank Brian Scheffler and Fanny Liu for the sequencing data. We are grateful to Bob Fjellstrom, Steven Brooks, Richard Jones and anonymous reviewers for critical reading and helpful comments of the manuscript. This research was in part funded by a grant from the USDA ARS.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.plantsci.2010.02.014.

References

[1] J.R. Harlan, The Living Fields: Our Agricultural Heritage, Cambridge University Press, New York, 1955, pp. 30–31.
[2] S.H. Hulbert, C.A. Webb, S.M. Smith, Q. Sun, Resistance gene complexes: evolution and utilization, Annu. Rev. Phytopathol. 39 (2001) 285–312.
[3] H.H. Flor, Current status of gene-for-gene concept, Annu. Rev. Phytopathol. 9 (1971) 275–296.
[4] S.H. Ou, Rice Diseases, Blast, Commonwealth Mycological Institute, Kew Surrey, UK, 1985, pp. 109–201.
[5] E. Ballini, J.B. Morel, G. Droc, A. Price, B. Courtois, J.L. Notteghem, D. Tharreau, A genome-wide meta-analysis of rice blast resistance genes and quantitative trait loci provides new insights into partial and complete resistance, Mol. Plant Microbe Interact. 21 (2008) 859–868.
[6] Z.X. Wang, M. Yano, U. Yamanouchi, M. Iwamoto, L. Monna, H. Hayasaka, Y. Katayose, T. Sasaki, The *Pib* gene for rice blast resistance belongs to the nucleotide binding and leucine-rich repeat class of plant disease resistance genes, Plant J. 19 (1999) 55–64.
[7] G.T. Bryan, K.S. Wu, L. Farrall, Y.L. Jia, H.P. Hershey, S.A. McAdams, K.N. Faulk, G.K. Donaldson, R. Tarchini, B. Valent, A single amino acid difference distinguishes resistant and susceptible alleles of the rice blast resistance gene *Pi-ta*, Plant Cell 12 (2000) 2033–2045.
[8] S.H. Qu, G.F. Liu, B. Zhou, M. Bellizzi, L.R. Zeng, L.Y. Dai, B. Han, G.L. Wang, The broad-spectrum blast resistance gene *Pi9* encodes a nucleotide-binding site-leucine-rich repeat protein and is a member of a multigene family in rice, Genetics 172 (2006) 1901–1914.
[9] B. Zhou, S.H. Qu, G.F. Liu, M. Dolan, H. Sakai, G.D. Lu, M. Bellizzi, G.L. Wang, The eight amino-acid differences within three leucine-rich repeats between *Pi2* and *Piz-t* resistance proteins determine the resistance specificity to *Magnaporthe grisea*, Mol. Plant Microbe Interact. 19 (2006) 1216–1228.
[10] X.W. Chen, J.J. Shang, D.X. Chen, C.L. Lei, Y. Zou, W.X. Zhai, G.Z. Liu, J.C. Xu, Z.Z. Ling, G. Cao, B.T. Ma, Y.P. Wang, X.F. Zhao, S.G. Li, L.H. Zhu, A B-lectin receptor kinase gene conferring rice blast resistance, Plant J. 46 (2006) 794–804.
[11] X.Q. Liu, F. Lin, L. Wang, Q.H. Pan, The in silico map-based cloning of *Pi36*, a rice coiled-coil-nucleotide-binding site-leucine-rich repeat gene that confers race-specific resistance to the blast fungus, Genetics 176 (2007) 2541–2549.
[12] F. Lin, S. Chen, Z.Q. Que, L. Wang, X.Q. Liu, Q.H. Pan, The blast resistance gene *Pi37* encodes a nucleotide binding site-leucine-rich repeat protein and is a member of a resistance gene cluster on rice chromosome 1, Genetics 177 (2007) 1871–1880.
[13] I. Ashikawa, N. Hayashi, H. Yamane, H. Kanamori, J. Wu, T. Matsumoto, K. Ono, M. Yano, Two adjacent nucleotide-binding site-leucine-rich repeat class genes are required to confer *Pikm*-specific rice blast resistance, Genetics 180 (2008) 2267–2276.
[14] S.K. Lee, M.Y. Song, Y.S. Seo, H.K. Kim, S. Ko, P.J. Cao, J.P. Suh, G. Yi, J.H. Roh, S. Lee, G. An, T.R. Hahn, G.L. Wang, P. Ronald, J.S. Jeon, Rice *Pi5*-mediated resistance to *Magnaporthe oryzae* requires the presence of two coiled-coil-nucleotide-binding-leucine-rich repeat genes, Genetics 181 (2009) 1627–1638.
[15] K. Hayashi, H. Yoshida, Refunctionalization of the ancient rice blast disease resistance gene *Pit* by the recruitment of a retrotransposon as a promoter, Plant J. 57 (2009) 413–425.
[16] J. Shang, Y. Tao, X. Chen, Y. Zou, C. Lei, J. Wang, X. Li, X. Zhao, M. Zhang, Z. Lu, J. Xu, Z. Cheng, J. Wan, L. Zhu, Identification of a new rice blast resistance gene, *Pid3*, by genomewide comparison of paired nucleotide-binding site-leucine-rich repeat genes and their pseudogene alleles between the two sequenced rice genomes, Genetics 182 (2009) 1303–1311.
[17] S. Fukuoka, N. Saka, H. Koga, K. Ono, T. Shimizu, K. Ebana, N. Hayashi, A. Takahashi, H. Hirochika, K. Okuno, M. Yano, Loss of function of a proline-containing protein confers durable disease resistance in rice, Science 325 (2009) 998–1001.
[18] T.R. Sharma, M.S. Madhav, B.K. Singh, P. Shanker, T.K. Jana, V. Dalal, A. Pandit, A. Singh, K. Gaikwad, H.C. Upreti, N.K. Singh, High-resolution mapping, cloning and molecular characterization of the *Pi-kh* gene of rice, which confers resistance to *Magnaporthe grisea*, Mol. Gen. Genom. 274 (2005) 569–578.
[19] X. Xu, N. Hayashi, C.T. Wang, H. Kato, T. Fujimura, S. Kawasaki, Efficient authentic fine mapping of the rice blast resistance gene *Pik-h* in the *Pik* cluster, using new *Pik-h*-differentiating isolates, Mol, Breed 22 (2008) 289–299.
[20] J.F. Bai, L.A. Pennill, J.C. Ning, S.W. Lee, J. Ramalingam, C.A. Webb, B.Y. Zhao, Q. Sun, J.C. Nelson, J.E. Leach, S.H. Hulbert, Diversity in nucleotide binding site-leucine-rich repeat genes in cereals, Genome Res. 12 (2002) 1871–1884.
[21] Y. Jia, G. Liu, S. Costanzo, S. Lee, Y. Dai, Current progress on genetic interactions of rice with rice blast and sheath blight fungi, Front. Agric. China 3 (2009) 231–239.
[22] K. Hayashi, N. Hashimoto, M. Daigen, I. Ashikawa, Development of PCR-based SNP markers for rice blast resistance genes at the *Piz* locus, Theor. Appl. Genet. 108 (2004) 1212–1220.
[23] D.H. Chen, M. Dela Viña, T. Inukai, D.J. Mackill, P.C. Ronald, R.J. Nelson, Molecular mapping of the blast resistance gene, *Pi44(t)*, in a line derived from a durably resistant rice cultivar, Theor. Appl. Genet. 98 (1999) 1046–1053.

[24] Z.H. Yu, D.J. Mackill, J.M. Bonman, S.R. McCouch, E. Guiderdoni, J.L. Notteghem, S.D. Tanksley, Molecular mapping of genes for resistance to rice blast (*Pyricularia grisea* Sacc), Theor. Appl. Genet. 93 (1996) 859–863.

[25] X. Sun, Y. Cao, Z. Yang, C. Xu, X. Li, S. Wang, Q. Zhang, *Xa26*, a gene conferring resistance to *Xanthomonas oryzae* pv. *oryzae* in rice, encodes an LRR receptor kinase-like protein, Plant J. 37 (2004) 517–527.

[26] C. Wang, M. Tan, X. Xu, G. Wen, D. Zhang, X. Lin, Localizing the bacterial blight resistance gene, *Xa22(t)*, to a 100-kilobase bacterial artificial chromosome, Phytopathology 93 (2003) 1258–1262.

[27] Y. Xiang, Y. Cao, C. Xu, X. Li, S. Wang, *Xa3*, conferring resistance for rice bacterial blight and encoding a receptor kinase-like protein, is the same as *Xa26*, Theor. Appl. Genet 113 (2006) 1347–1355.

[28] S. Kiyosawa, Genetics of blast resistance, in: International Rice Research Institute (Ed.), Rice Breeding, International Rice Research Institute, Los Banos, Manila, Philippines, 1972, pp. 203–225.

[29] S.R. McCouch, R.J. Nelson, J. Tohme, R.S. Zeigler, Mapping of blast resistance genes in rice, in: R.S. Zeigler, S.A. Leong, P.S. Teng (Eds.), Rice Blast Disease, CABI, Wallingford, Oxon, UK, 1994, pp. 167–186.

[30] K. Yoshida, H. Saitoh, S. Fujisawa, H. Kanzaki, H. Matsumura, K. Yoshida, Y. Tosa, I. Chuma, Y. Takano, J. Win, S. Kamoun, R. Terauchi, Association genetics reveals three novel avirulence genes from the rice blast fungal pathogen *Magnaporthe oryzae*, Plant Cell 21 (2009) 1573–1591.

[31] I. Ashikawa, J. Wu, T. Matsumoto, R. Ishikawa, Haplotype diversity and molecular evolution of the rice *Pikm* locus for blast resistance, J. Gen. Plant Pathol. (2009) 1–6.

[32] K.A.K. Moldenhauer, F.N. Lee, R.J. Norman, R.S. Helms, B.R. Wells, R.H. Dilday, P.C. Rohman, M.A. Marchetti, Registration of Katy rice, Crop Sci. 30 (1990) 747–748.

[33] B. Valent, L. Farrall, F.G. Chumley, *Magnaporthe grisea* genes for pathogenicity and virulence identified through a series of backcrosses, Genetics 127 (1991) 87–101.

[34] S. Rozen, H. Skaletsky, Primer3 on the WWW for general users and for biologist programmers, Methods Mol. Biol. 132 (2000) 365–386.

[35] J. Sambrook, E.F. Fritsch, T.M. Maniatis, Molecular Cloning: A Laboratory Manual, 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989.

[36] K. Tamura, J. Dudley, M. Nei, S. Kumar, MEGA4: Molecular evolutionary genetics analysis (MEGA) software version 4.0, Mol. Biol. Evol. 24 (2007) 1596–1599.

[37] M. Gruber, J. Soding, A.N. Lupas, Comparative analysis of coiled-coil prediction methods, J. Struct. Biol. 155 (2006) 140–145.

[38] P. Rice, I. Longden, A. Bleasby, EMBOSS: the European molecular biology open software suite, Trends Genet. 16 (2000) 276–277.

[39] T.A. Hall, BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT, Nucl. Acids Symp. Ser. 41 (1999) 95–98.

[40] A.C. Syvanen, From gels to chips: "minisequencing" primer extension for analysis of point mutations and single nucleotide polymorphisms, Hum. Mutat. 13 (1999) 1–10.

[41] P. Burkhard, J. Stetefeld, S.V. Strelkov, Coiled coils: a highly versatile protein folding motif, Trends Cell Biol. 11 (2001) 82–88.

[42] J.M. Salmeron, G.E.D. Oldroyd, C.M.T. Rommens, S.R. Scofield, H.S. Kim, D.T. Lavelle, D. Dahlbeck, B.J. Staskawicz, Tomato *Prf* is a member of the leucine-rich repeat class of plant disease resistance genes and lies embedded within the *Pto* kinase gene cluster, Cell 86 (1996) 123–133.

[43] D.A. Jones, J.D.G. Jones, The role of leucine-rich repeat proteins in plant defences, Adv. Bot. Res. Adv. Plant Pathol. 24 (1996) 89–167.

[44] Y. Jia, F.N. Lee, A. McClung, Determination of resistance spectra of the *Pi-ta* and *Pi-k* genes to U.S. races of *Magnaporthe oryzae* causing rice blast in a recombinant inbred line population, Plant Dis. 93 (2009) 639–644.

[45] Y. Jia, R. Martin, Identification of a new locus, *Ptr(t)*, required for rice blast resistance gene *Pi-ta*-mediated resistance, Mol. Plant Microbe Interact. 21 (2008) 396–403.

[46] S. Kiyosawa, Inheritance of resistance of rice varieties to a Philippine fungus strain of *Pyricularia oryzae*, Jpn. J. Breed. 19 (1969) 61–73.

[47] S. Kiyosawa, Inheritance of blast resistance in some Chinese rice varieties and their derivatives, Jpn. J. Breed. 18 (1968) 193–204.

[48] S. Kiyosawa, Identification of blast resistance genes in some rice varieties, Jpn. J. Breed. 28 (1978) 287–296.

[49] Y. Yamasaki, S. Kiyosawa, Studies on inheritance of resistance of rice varieties to blast, 1. Inheritance of Japanese varieties to several strains of the fungus, Bull. Natl. Inst. Agric. Sci. D14 (1966) 39–69.

[50] S. Kiyosawa, H. Ikehashi, H. Kato, Z.Z. Ling, Pathogenicity tests of Philippine isolates of blast fungus using two sets of rice varieties, Jpn. J. Breed. 31 (1981) 367–376.